**Commentaries on**
*The Space Environment: Implications for Spacecraft Design*

by
Alan C. Tribble

Princeton University Press, 1995

# Chapter 1
# Introduction

## 1.4 The Earth's Magnetic Field

The Earth's magnetic field is very nearly a dipole field with vector components in a spherical polar coordinate system $(r, \theta, \phi)$given by eq. 1.2 on p.9.:

$$B_r = \frac{\partial V}{\partial r} = -\frac{2M}{r^3}\cos\theta$$
$$B_\theta = \frac{1}{r}\frac{\partial V}{\partial \theta} = -\frac{M}{r^3}\sin\theta$$
$$B_\phi = \frac{1}{r\sin\theta}\frac{\partial V}{\partial \phi} = 0.$$

The geometrical relationships between the magnetic moment **M**, the radius vector **r** and the components of the magnetic field vector **B** are shown in the figure.

The geomagnetic field does not exist in isolation but is immersed in the solar magnetic field. The geomagnetic field falls off as $1/r^3$ eventually becoming weaker than the surrounding solar magnetic field. The locus of points at which the geomagnetic field matches the solar magnetic field in intensity is called the *magnetopause*. Within the magnetopause, the earth's magnetic field is the stronger; outside of the magnetopause, the solar magnetic field is the stronger.

The geomagnetic field closely resembles an ideal dipole field close to the earth; i.e. within ten earth radii ($< 10R_E$). Beyond $10R_E$, the shape of the field is modified substantially by the action of the *solar wind*. The solar wind is a stream of energetic charged particles, mostly protons and electrons, ejected from the sun and blowing outward toward interstellar space. This stream of charged particles constitutes an electric current flowing past the earth which distorts the geomagnetic field. Windward of the earth (toward the sun) the magnetic field lines are compressed inward toward the surface of the earth; leeward, the magnetic field lines are stretched out into a long *magnetotail* extending as much as $100R_E$ into the solar wake.

The geomagnetic field is believed to be caused by electric currents deep within the body of the earth. Heat generated by radioactive decay in the solid core drives convection cells in the liquid mantle. The material of the mantle is metallic and therefore conducting. As the convection cells move the conducting material of the mantle through local magnetic fields, electric currents are induced which, in turn, produce their own magnetic fields. These magnetic fields interact with the surrounding convection cells to induce still more electric currents. The induced currents and the inducing magnetic fields lock in a positive feedback loop which produces the geomagnetic field as steady state output. This action is analogous to that of a self-exciting direct-current generator.

## 1.5 The Solar-Planetary Relationship

The Sun is a common, ordinary star of the type described by astronomers as a yellow dwarf of spectral type G2 V. It has a surface temperature of 5770 K, a surface gravity of $2.74 \times 10^2$ m s$^{-2}$, mass of $1.99 \times 10^{30}$ kg, and radius $6.96 \times 10^8$ m. It is a late-type main-sequence dwarf, believed to be five billion years old and expected to shine more or less as it is now for another five billion years.

The solar constant is defined as the amount of energy originating in the sun deposited at the top of the earth's atmosphere. It represents the spectral luminance graphed in Figure 2 integrated over all wavelengths. The solar constant averages $1371 \pm 5$ W/m$^2$ but varies with the seasons due to the ellipticity of the earth's orbit. This energy is overwhelmingly electromagnetic radiation but does contain a very small contribution in the form of particulate radiation from the solar wind. The luminosity of the sun, defined as its total power output, is $3.90 \times 10^{33}$ W.

The enormous energy emitted by the sun originates deep in the core of the sun in thermonuclear reactions which convert hydrogen into helium. In the process, a small amount of mass in converted into energy. Most of the energy is produced in the proton-proton cycle as follows:

$$\begin{array}{rcll} {}_1\mathrm{H}^1 +_1 \mathrm{H}^1 & \longrightarrow & {}_1\mathrm{H}^2 + e^+ + \nu & 1.44\ \mathrm{MeV} \\ {}_1\mathrm{H}^2 +_1 \mathrm{H}^1 & \longrightarrow & {}_2\mathrm{He}^3 + \gamma & 5.49\ \mathrm{MeV} \\ {}_2\mathrm{He}^3 +_2 \mathrm{He}^3 & \longrightarrow & {}_2\mathrm{He}^4 +_1 \mathrm{H}^1 +_1 \mathrm{H}^1 & 12.85\mathrm{MeV} \end{array}$$

${}_1\mathrm{H}^2$ is deuterium, an isotope of hydrogen with one proton and one neutron in the nucleus. $e^+$ is a positron (an anti-electron or positive electron), $\nu$ is a neutrino, and $\gamma$ is a gamma-ray photon. The values at the right give the amount of energy released in each reaction. Note that the first reaction must take place twice in order to produce one helium nucleus; the total energy liberated per helium nucleus is 26.2 MeV. The neutrinos escape without any further interaction with the bulk of the sun. The 0.26 MeV they carry contributes nothing to the thermal balance of either the sun or of the earth and is *not* included in the values of the solar luminosity or of solar constant quoted above.

Helium production can also take place through a different mechanism called the carbon-nitrogen-oxygen cycle:

$$\begin{array}{rcll} {}_6\mathrm{C}^{12} +_1 \mathrm{H}^1 & \longrightarrow & {}_7\mathrm{N}^{13} + \gamma & 1.94\ \mathrm{MeV} \\ {}_7\mathrm{N}^{13} & \longrightarrow & {}_6\mathrm{C}^{13} + e^+ + \nu & 2.22\ \mathrm{MeV} \\ {}_6\mathrm{C}^{13} +_1 \mathrm{H}^1 & \longrightarrow & {}_7\mathrm{N}^{14} + \gamma & 7.55\ \mathrm{MeV} \\ {}_7\mathrm{N}^{14} +_1 \mathrm{H}^1 & \longrightarrow & {}_8\mathrm{O}^{15} + \gamma & 7.29\ \mathrm{MeV} \\ {}_8\mathrm{O}^{15} & \longrightarrow & {}_7\mathrm{N}^{15} + e^+ + \nu & 2.76\ \mathrm{MeV} \\ {}_7\mathrm{N}^{15}_1\mathrm{H}^1 & \longrightarrow & {}_6\mathrm{C}^{12} +_1 \mathrm{He}^4 & 4.97\ \mathrm{MeV} \end{array}$$

The decay of the ${}_7\mathrm{N}^{13}$ nucleus produces a 0.7 MeV neutrino. The decay of the ${}_8\mathrm{O}^{15}$ nucleus produces a 1.00 MeV neutrino. As before, this energy in not deposited in the bulk of the sun so that this cycle contributes 25 Mev to the thermal balance of the star, slightly less than that of the proton-proton cycle.

These thermonuclear reactions take place deep within the core of the sun where the temperature exceeds 13 million K and reaches 15 million K at the center. Temperature falls with radius and thermonuclear activity ceases at 10 million K. This radius marks the boundary between the solar core and the radiation zone. Energy leaves the core in the form of gamma radiation proceeding outward toward the surface in a random walk of absorption, reemission and scattering. As it moves outward, it cools to match the temperature of the surrounding hydrogen plasma, becoming first x-radiation and then ultraviolet radiation. When the temperature has fallen to 1.5 million K the hydrogen plasma becomes opaque to ultraviolet radiation and the radius at which this happens is the outer boundary of the radiation zone.

At this point, the energy flows outward toward the surface through convection. Hot hydrogen plasma rises from the boundary of the radiation zone to the surface where it loses energy through radiation out into space, cools and sinks. When the plasma returns to the radiation zone it is reheated and rises again.

The outer surface of the convection zone constitutes the visible surface of the sun called the photosphere. Negative hydrogen ions within the plasma make the photosphere opaque to visible radiation. The visible surface of the sun represents the radius at which the temperature is too cool for negative hydrogen to exist. The drop in concentration of negative hydrogen is rapid enough to give the sun a sharp, well-defined edge. The hydrogen above the photosphere is transparent to visible light and invisible from earth. From this point on the radiant energy emitted by the sun flows unimpeded outward into the solar system and beyond into interstellar space.

The presence of the convection zone beneath the photosphere is revealed in high-resolution photographs of the photosphere. The photographs reveal a granulated surface where the hot bright ascending gas in the center of a granulation is surrounded by a dark ring of cool descending gas. Scattered among the granules are large dark regions known as sunspots. Each sunspot is a region of concentrated magnetic flux thousands of times stronger than the magnetic flux found in nearby quieter regions. The trapped magnetic field acts as a cooling mechanism that dampens turbulent flow within the spot and forces hot bright gas to surface elsewhere. The gas within sunspots has a temperature of 3800 K; still hot enough to glow brighter than the moon, but cool enough to look dark in comparison with the hotter surrounding photosphere. Small sunspots run a few thousand miles in diamerer, large enough to engulf the earth. The largest on record exceeded 7 billion square miles.

Above the photosphere lies the first layer of the solar atmosphere, the chromosphere. This is a thin but violently active region driven by the turbulent convection cells in the photosphere below. Outside the chromosphere lies the corona. The chromosphere is filled hot jets of gas, called spicules, thrown up by bursting granules. Each spicule weighs a million tons, glows at 100,000 K, and rises to heights of five thousand miles or so with an initial velocity of 10 to 15 miles per second.

Prominences are huge streamers of hot gas erupting from the chromosphere into the corona. They are tied to the spots of greatest magnetic field strength in the photosphere and are formed by hot ionized gas tied to a magnetic field line. As the magnetic field twists and turns in the turbulence of the photosphere, the material in the prominence is whipped violently back and forth, occasionally breaking off and streaming into space. Small prominences remain close to the photosphere forming closed loops that give the sun "the appearance of a closely knit sweater." (Friedman, 1986, p. 81)

Plages are regions of strong magnetic activity that glow distinctly brighter than their surroundings.

Temperature rises with altitude in the chromosphere and corona, reaching temperatures of one million K or higher in the outermost regions. Because of the extremely low density of this high altitude superheated gas, the glow from it can only be seen when the body of the sun is blocked by the moon in a total solar eclipse. At one million K, the peak of the radiation curve is in the x-ray region and the corona is responsible for what little x-radiation reaches the earth from the sun. According to one hypothesis, the anomalously high temperature of the corona is due to turbulent activity in the photosphere. This generates powerful acoustic waves in the chromosphere that increase in velocity as they climb into the rarified gas of the corona, eventually becoming supersonic and dissipating their energy as heat.

### 1.5.1 The 11-Year Solar Cycle

The number and size of sunspots varies on a cycle of approximately 11 years. The shortest cycle on record lasted seven years, the longest 17 years. Each cycle begins with a quiet sun relatively free of sunspots. Early in the cycle, sunspots begin to appear at 30° N and 30°S solar latitude. As the cycle progresses, new spots appear closer to the equator and in more profusion. The last spots formed are found at 5°N and S. Frequently, the new cycle begins at 30° N and S before the spots from the old cycle have completely faded. Sunspots in the northern hemisphere have opposite magnetic polarity from those in the southern hemisphere, and the polarity reverses with each new 11-year cycle. The underlying mechanism therefore has a 22-year cycle. What that underlying mechanism may be is a matter of speculation.

The cycle has displayed a great deal of historical variability. The most extreme variation occurred from 1660 to 1710 when no sunspots at all would be seen for years at a time. It was considered a remarkable occurrence in 1705 when two sunspots were seen on the face of the sun at once. Perhaps coincidentally (perhaps not) this was a time of severe cold weather in Europe and North America. During this period, the Thames River (England) and the English Channel routinely froze over in wintertime thick enough to support skaters.

Carbon-14 studies in the 5000-year old bristlecone pines of California's White Mountains allow the solar cycle to be followed back to times before sunspot records were kept. These and other isotope studies indicate the existence of the Sporer minimum during which the Arctic icepack expanded far enough south to push the Norsemen out of Greenland and snow accumulated on high mountain peaks in Ethiopia. Many suggestive coincidences like these between the climatic record and the solar cycle argue for a link between solar cycle activity and terrestrial weather, but the exact relationship, if any, is poorly understood. What is clear from the carbon-14 studies is that Maunder-like minima have not been rare over the past few millenia. The type of active sun we see today may in fact be the exception rather than the norm.

### 1.5.2 The Solar Wind and Solar Flares

Unlike the terrestrial atmosphere, the solar atmosphere is not gravitationally bound. At a temperature of one million K, the particles making up the corona have a velocity well in excess of solar escape velocity. High altitude rocket and satellite observations of the late 1950's and early 1960's revealed the presence of tightly meshed magnetic loops at the base of the corona which hold the coronal plasma to the sun. At the solar poles, however, and at sites of intense sunspot activity, the magnetic field lines extend deep into

interplanetary space before closing back on the sun. In these regions, the ionized gas of the corona is free to stream out into space and escape from the sun. The sun is literally boiling itself off into interstellar space. This stream of material outward from the surface of the sun is called the solar wind.

The material of the solar wind follows the magnetic field lines away from the sun but the magnetic field strength weakens with distance and eventually has negligible effect on the flow pattern. Near the sun where the field is strong the solar wind is tightly bound to the field and corotates with the sun. Farther out the influence of the field diminishes and the wind lags behind the field lines. This is the reason behind the flow pattern displayed in Fig. 1.9 on page 17.

Small solar flares are frequent and release the energy equivalent of one million hydrogen bombs. The largest and rarest flares are the equivalent of one billion hydrogen bombs detonated in the span of a few hours.

### 1.5.3 Geomagnetic Storms

The solar wind travels outward at 300 to 750 km/s and has a temperature on the order of $10^5$ K. The speed of sound at this temperature is given by

$$v_s = \sqrt{kT/m_p} \sim 30 \quad \text{km/s}$$

so that the Mach number of the solar wind is $\sim$ 10-25. The earth and its magnetosphere constitute a nearly stationary obstacle in this hypersonic flow, generating a bow shock.

The bow shock lies 2-3 $R_E$ outside of the magnetopause. The turbulent layer immediately behind the bow shock and in front of the magnetopause is called the magnetosheath. Electrons and protons is the magnetosheath can attain energies several thousand times that of the incoming solar wind. The magnetosphere cusps at the poles as the magnetic field lines funnel down to the magnetic poles. Leeward, the field lines are stretched out nearly parallel to the solar wind but oppositely directed above and below the magnetic equator. Ampere's Law

$$\oint \mathbf{B} \cdot d\mathbf{s} = \mu I_{\text{enc}}$$

when applied to the loop shown in Figure 3 shows that there must be a sheet of current running clockwise along the magnetic equator as seen from someone looking down on the north magnetic pole. This current forms a plasma sheet separating the oppositely directed magnetic field lines.

Gusts in the solar wind move this bow shock toward and away from the earth at distances between 77,000 to 83,000 kilometers (12-13 $R_E$) at speeds of 10 to 200 km/s. The resulting change in the size and shape of the magnetosphere is called a geomagnetic storm. Magnetic field strength changes everywhere within the magnetospere as the flux lines wiggle back and forth under the pressure of the solar wind. The changing flux induces electromotive force (emf) in conducting loops via Faraday's Law:

$$\oint \mathbf{E} \cdot d\mathbf{s} = -\frac{d}{dt} \oint \mathbf{B} \cdot d\mathbf{A}$$

If the loops enclose a sufficiently large area or have a sufficiently large number of flux linkages the resulting emf will be large enough to have noticeable effects. Telephone systems and electric power grids are fundamentally huge networks of conducting loops capable of building up very strong emfs in the presence of geomagnetic storms.

In August-September 1859, telegraph lines out of Boston intermittently stopped working during severe solar disturbances. At other times the lines worked fine with the batteries disconnected, powered instead by the emf induced in the lines by the geomagnetic storms. Telegraph operators reported that the induced line voltages correlated with the appearance of aurorae over the Boston area. On 24 March 1970, 78 percent of long distance calls from Minneapolis were lost to a geomagnetic storm. On 10 February 1958 geomagnetically induced emf in a Newfoundland-to-Scotland transatlantic cable reached 2650 volts. Grids and systems in Canada are very susceptible to geomagnetic disruption because of the proximity of the north geomagnetic pole off the western coast of Greenland: geomagnetic disturbances on 9 February 1958 tripped circuit breakers

in Ontario, blacking out Toronto; a solar flare in 1972 destroyed a 230-kilovolt power transformer in British Columbia. Less dramatic but more common are disruptions and interference in communications. (Friedman, 1986, pp. 183-4.)

Solar protons released by solar flares take one or two days to reach the earth. When they do so, the geomagnetic field shunts them into the atmosphere over the magnetic poles producing enhanced aurorae. More energetic flares produce 10-500 million electron volt protons with energies comparable to galactic cosmic rays. These protons are too energetic to be deflected by the geomagnetic field and smash into the upper atmosphere leading to secondary cosmic rays showered over a wide area.

### 1.8 References

1. Friedman, H. 1986. Sun and Earth. New York: Scientific American Library.

# Chapter 2
# The Vacuum Environment

## 2.1 Overview

The presence of an atmosphere here on Earth provides a number of benefits such as pressure, oxygen for respiration, humidity, thermal regulation, micrometeoroid protection, and radiation shielding. None of these are naturally available in the near-vacuum of space and when needed must be included in the spacecraft design.

The composition of the atmosphere at sea level is 79% nitrogen ($N_2$) and 20% oxygen ($O_2$) with the remaining 1% a mixture of trace gases (Ar, $CO_2$, etc). To an excellent approximation, the atmosphere obeys the ideal gas equation of state

$$P_0 = n_0 kT$$

where $P_0$ is the total pressure, $n_0$ is the number density of atoms or molecules, $k$ is Boltzmann's constant, and $T$ is the absolute temperature (measured in Kelvins). Since the number density $n$ is the sum of the number densities of the constituent gases,

$$n_0 = n_{N_2} + n_{O_2} + n_{Ar} + n_{CO_2} + \ldots$$

the total pressure is just the sum of the *partial pressures* of the constituent gases

$$P_0 = P_{N_2} + P_{O_2} + P_{Ar} + P_{CO_2} + \ldots$$

This is Dalton's Law of Partial Pressures. For each constituent, the ratio of partial pressure to total pressure equals the ratio of the corresponding number densities:

$$\frac{P_X}{P_0} = \frac{n_X}{n_0}$$

so that $P_{N_2} = 0.79P_0$; $P_{O_2} = 0.2P_0$; etc.

Normal human respiration at sea level takes place at a partial pressure of oxygen ($PO_2$) of $0.2P_0 =$ 20 kiloPascals (kPa)= 3 lb/in$^2$ (psia). As atmospheric pressure drops exponentially with increasing altitude, so does $PO_2$. Figure 2.1 shows the range of barometric pressure tolerated by humans when $PO_2$ is maintained within safe limits. Note that lower total ambient pressures are acceptable if the percentage of oxygen is increased. The lungs always contain water vapor at a partial pressure of 0.9 psia (the vapor pressure of water at body temperature) and $CO_2$ at 0.7 to 0.9 psia (the result of normal respiration). Oxygen saturation of arterial blood can be maintained at a $PO_2$ of 1.9 psia, so that normal human activity can continue at a total barometric pressure of 3.6 to 3.8 psia if the ambient atmosphere consists of pure oxygen; where pure oxygen has been used in US aviation and space activities, the practice has been to maintain 5 psia ambient pressure. Higher ambient pressures that maintain $PO_2$ levels of 3.0 psia are required if the oxygen is diluted with other gases.

At $PO_2$ levels less than 2.0 psia, the effects of decreased oxygen intake (*hypoxia*) become apparent. Exposure to $PO_2$ levels of 1.2-2 psia lead to *acute altitude sickness*, the symptoms of which are shortness of breath, headaches, insomnia, impaired concentration, decrease in ability to perform complex tasks, nausea, and vomiting. Continued exposure to $PO_2$ at these levels will often be accompanied by some degree of acclimatization, but the ability to perform physical activity will be noticeably impaired and continued exposure will exact a physiological toll. $PO_2$ levels below 1.2 psia eventually lead to loss of consciousness. Recovery is possible if recompression (restoration of normal $PO_2$) takes place quickly enough, but the required recompression time becomes shorter as $PO_2$ decreases. Complete loss of $PO_2$ is referred to as *anoxia*, and

leads to loss of consciousness in 9 to 11 seconds. Shortly thereafter, water vapor begins to form in the veins, heart rate falls rapidly, and with it, blood pressure in the arteries. After one minute pressure in the veins exceeds pressure in the arteries and blood flow ceases. Recompression must take place within 90 seconds if survival is to be assured.

No American astronauts have been lost to hypoxia but the same is not true for the Russian space program. On 1 November 1962, P. I. Dolgov died during high altitude space suit tests outside a ballon gondola at 25,485 ft. While exiting the gondola, he broke his suit helmet and died from acute hypoxia. On 30 June 1971, G. T. Dobrovolskiy, V. N. Volgorov, and V. I. Patsayev died of hypoxia when their Soyuz 11 descent module lost pressure during re-entry. The cosmonaut closest to the valve which failed managed to unfasten his seatbelt, but the other two died before they had time to respond to the emergency.

Oxygen supply is not the only source of concern. Even if $PO_2$ remains high enough to insure adequate oxygenation of the blood, a large drop in ambient pressure can have detrimental effects. Explosive decompression rapidly lowers the external pressure on the human body, but trapped air in the lungs, gastrointestinal tract, middle ear, and sinus cavities remains at the former pressure and now pushes outward. Tissue distension and rupture are possible. Significant rupture of the pulmonary alveoli can lead to shock, gas embolism, and death. The pressure differential damage threshold for healthy human lungs is unknown but is believed to be about 1.5 psia. Less rapid decompression allows this trapped air to escape through the available orifices without any significant pressure differential developing.

Continued exposure to low ambient pressures with $PO_2$ maintained at adequate levels can lead to *altitude decompression sickness*, the result of gas bubbles forming in the body. Nitrogen dissolved in the fluid and tissues of the human body is held in solution by the ambient pressure. When ambient pressure falls, fluid and tissue become supersaturated with this substance and evolution as a free gas is possible. Pre-existing bubbles in body tissues above a certain critical size will collect dissolved nitrogen and grow. The gas bubbles then exert pressure on the surrounding tissues, distorting and distending them, and producing pain. The onset of altitude decompression sickness is hastened by vigorous physical activity which accelerates the formation of bubbles capable of growing under supersaturated conditions.

Because of the large variation in possible locations of the bubbles, symptoms of altitude decompression sickness vary greatly. The most commom symptom is pain in the joints, known as the bends. Less commom symptoms include itching, skin rash, edema (fluid collecting beneath the skin), and change in skin color. Severe forms of altitude decompression sickness occur when bubbles collect in the pulmocirculatory and neurocirculatory systems. Bubbles in the pulmocirculatory system are particularly dangerous. Victims experience a deep pain in the chest on inhalation, a dry cough, and choking fits. Severe cases of coughing and choking lead to loss of consciousness and collapse. Bubbles in the neurocirculatory system are extremely dangerous. The symptoms are general malaise, headaches, and symptoms typical of local injury to the brain and spinal cord.

Treatment for altitude decompression sickness involves recompression of the victim as early as possible after the onset of symptoms. The sickness can be prevented by washout of dissolved nitrogen either by slowly reducing ambient pressure or prebreathing pure oxygen. The amount of time which must be devoted to prebreathing pure oxygen, typically a few hours, increases with decreasing ambient pressure.

Partial pressures of oxygen well above sea level values can lead to oxygen toxicity: hyperoxia.

Excessive levels of carbon dioxide can be dangerous even if $PO_2$ is more than adequate to support respiration. Partial pressures of $CO_2$ above 0.5 psia cause hypercapnia. At sea level this would correspond to a concentration of 3.3%. The normal concentration of $CO_2$ at sea level is typically 0.2%.

Manned spacecraft must be designed to provide an ambient atmosphere capable of preventing both altitude sickness and altitude decompression sickness. The easiest way of doing this is to simply supply an atmosphere identical to sea level ambient in composition and pressure: 80% $N_2$ and 20% $O_2$ at 14.7 psia. However, the high partial pressure of nitrogen requires extended prebreathing sessions to avoid altitude decompression sickness during extravehicular activity (EVA); containing 14.7 psia requires thick and strong cabin walls with an accompanying weight penalty; gas leakage to the external vacuum is accelerated, particularly during opening and closing of airlocks; and the high gas density requires correspondingly strong fans and pumps with high power demands for cabin ventilation.

This is one extreme. The other is to suppply 100% $O_2$ at 5 psia. This reduces the difficulties outlined above, but creates a significant fire hazard. Such was the practice in the US manned space program before

the Apollo 1 command module fire on 27 January 1967. The Russian space program originally considered a similar atmosphere but abandoned it after a February 1961 fire in an altitude chamber filled with pure oxygen at intermediate pressure killed a cosmonaut candidate named Bondarenko. All subsequent Russian manned flights, and all US Space Shuttle flights have used sea level equivalent pressures and compositions to avoid similar tragedies.

As originally conceived, the Apollo command module would have contained pure oxygen at 14 psia during countdown, with the pressure dropping during the ascent phase to stabilize at 5 psia on orbit. The countdown rehearsal which ended in the deaths of Virgil Grissom, Edward White, and Richard Chaffee was conducted in this atmosphere. An electrical spark under the command pilot's seat ignited packing material which had been left in the command module. Fed by pure oxygen at sea level pressure, the fire spread rapidly to consume most of the flammable material within the command module in a matter of minutes. Temperature in the command module exceeded 1500 ° F. in less than four minutes. The astronauts were fully suited but temperatures inside the spacesuits exceeded 800 ° F. Pressure increased with temperature, reaching a maximum of 30 psia and rupturing the command module shell. All three astronauts died in a matter of minutes from smoke inhalation and heat.

As a result of this tragedy, all subsequent Apollo flights used sea level equivalent atmospheres during countdown, with this atmosphere gradually replaced by pure oxygen at 5 psia upon achieving orbit. This pure oxygen atmosphere was maintained throughout the mission until it was replaced by sea level equivalent again upon re-entry. This created problems during the Apollo-Soyuz Test Project. The Apollo command module with pure oxygen at 5 psia could not dock directly with the Soyuz module containing 80/20 $N_2/O_2$ mix at 14 psia without exposing the Soyuz crew to rapid decompression and the risk of the bends. The two spacecraft had to be linked with a special tunnel containing an atmosphere of intermediate composition and pressure.

All US Space Shuttle flights have used a sea level equivalent atmosphere.

### 2.2.3 Effects of Contamination

The derivation of equation (2.19) in Tribble (p. 36) is fairly straightforward. The uncontaminated surface has absorbtance $\alpha_s(\lambda)$. Energy balance demands that

$$S(\lambda) = \alpha_s(\lambda)S(\lambda) + R_s(\lambda)S(\lambda)$$

or

$$\alpha_s(\lambda) + R_s(\lambda) = 1.$$

The absorptance of the surface with the contaminant layer is by definition the total radiation absorbed divided by the total radiation incident. The total radiation absorbed is found by following the radiation step-by-step through the contamination layer.

Consider first the initial transmission of radiation through the contaminant. For thin layers, the absorption of radiation as a fraction of the incident radiation will be a linear function of the thickness of the layer:

$$\frac{\Delta E}{E} = \alpha_c(\lambda)\,\Delta x.$$

If, for example, a 1-mm layer absorbs 10% of the radiation falling on it, then a 2-mm layer will absorb 20%, and so on. For finite layers, this law is only an approximation, but one which becomes better as the layer becomes thinner, so that in the appropriate limit, the relation becomes an exact differential

$$\frac{dE}{E} = -\alpha_c(\lambda)\,dx$$

where the minus sign has been added to indicate that absorption takes place and the transmitted energy decreases with depth.

The energy remaining $E(x)$ can be found by integrating the equation.

$$\begin{aligned}\int_{E_0}^{E} \frac{dE}{E} &= \int_0^x -\alpha_c(\lambda)\,dx \\ \ln\left(\frac{E}{E_0}\right) &= -\alpha_c(\lambda)x \\ E &= E_0 e^{-\alpha_c(\lambda)x} = E(x).\end{aligned}$$

$E(x)$ is the energy transmitted to the depth $x$, and $E_0$ is the energy incident on the contaminant layer. For our scenario, $S(\lambda)$ is the incident radiation, $x$ is the thickness of the contaminant layer, and

$$S(\lambda)e^{-\alpha_c(\lambda)x}$$

is the amount of radiation transmitted to the spacecraft surface.

Of the radiation transmitted to the spacecraft surface, a fraction $R_s(\lambda)$ will be reflected and a fraction $\alpha_s(\lambda)$ absorbed. The total amount reflected will be

$$R_s(\lambda)e^{-\alpha_c(\lambda)x}S(\lambda)$$

and this radiation must traverse the contaminant layer again. The amount of this radiation transmitted will be

$$\left\{R_s(\lambda)e^{-\alpha_c(\lambda)x}S(\lambda)\right\}e^{-\alpha_c(\lambda)x} = R_s(\lambda)e^{-2\alpha_c(\lambda)xS(\lambda)}.$$

This radiation will escape from the spacecraft and represents the radiation *reflected*. The radiation *absorbed* is

$$S(\lambda) - R_s(\lambda)e^{-2\alpha_c(\lambda)x}S(\lambda) = \left\{1 - R_s(\lambda)e^{-2\alpha_c(\lambda)x}\right\}S(\lambda).$$

The *total radiation absorbed* is found by integrating over the entire spectrum:

$$\int_0^\infty \left\{1 - R_s(\lambda)e^{-2\alpha_c(\lambda)x}\right\}S(\lambda)\,d\lambda.$$

The total absorptance is then

$$\alpha_s(\lambda) = \frac{\int_0^\infty \left\{1 - R_s(\lambda)e^{-2\alpha_c(\lambda)x}\right\}S(\lambda)\,d\lambda}{\int_0^\infty S(\lambda)\,d\lambda}. \qquad \text{(2.19 p. 36)}$$

The situation is similar for solar cells. The power delivered by a solar cell is some fraction of the incident power

$$I_s(\lambda)S(\lambda)e^{-\alpha_c(\lambda)x}$$

where $I_s(\lambda)$ is the *solar cell response*, also known as the *quantum efficiency*. Here, the incident solar power $S(\lambda)$ has been degraded by absorption in the contaminant layer to

$$S(\lambda)e^{-\alpha_c(\lambda)x}.$$

Of that power, only the fraction $I_s(\lambda)$ becomes useful electrical power.

**Chapter 4**
**The Plasma Environment**

4.4.1.1 Dielectric Breakdown

Dielectrics are nonconducting materials. When inserted between the two plates of a capacitor completely filling the gap, their effect is to increse the capacitance by a multiplying factor $K$ called the dielectric constant:

$$C = KC_0.$$

For an isolated capacitor, the charge must remain constant so the net effect of the presence of a dielectric is to increase the capacitor voltage relative to the vacuum value:

$$Q = \frac{C}{V} = \frac{KC_0}{KV_0}$$

where $C_0$ and $V_0$ are the vacuum values. If the capacitor is connected to a constant-voltage source, the net effect is to increase the charge residing on the plates of the capacitor:

$$V = \frac{C}{Q} = \frac{KC_0}{KQ_0}$$

where $Q_0$ is the vacuum value.

Both effects are due to the electric polarization of the dielectric molecules. Every atom is composed of a positive nuclear core surrounded by an electron cloud. In the presence of an external field such as the electric field that exists between the two plates of a capacitor, the positive nuclei are drawn in the direction of the field while the negative electrons are drawn opposite the direction of the field. The resulting charge separation creates an electric field within the atom that exactly balances the external field and maintains equilibrium. The internal field, however, is oppositely directed to the external field, partially canceling it and overall reducing the net field within the dielectric and thus between the plates. Since the potential difference is

$$\Delta V = -\int \mathbf{E}_{\text{net}} \cdot d\mathbf{s}$$

the weaker field leads to a smaller voltage for a given charge and, in turn, a greater capacitance. For a constant voltage source, a greater charge on the capacitor plates is required to provide an external field strong enough to provide the required potential difference after partial cancelation by the dielectric. Since the dielectric is a nonconductor, the electric field within the dielectric, albeit weaker than the external field, is not zero.

For a given $K$, $C$, and $V$, the net potential is given by

$$V_{net} = \frac{Q}{KC}$$

while the vacuum potential is given by

$$V_{vac} = \frac{Q}{C}.$$

The contribution to the potential due to the internal dielectric field is then

$$V_{int} = V_{net} - V_{vac} = \frac{Q}{KC} - \frac{Q}{C} = \frac{Q}{C}\left(\frac{1}{K} - 1\right) = \left(\frac{1-K}{K}\right) V_{vac}$$

so that the internal electric field itself is

$$E_{int} = \frac{V_{int}}{d} = \left(\frac{1-K}{K}\right)\frac{V_{vac}}{d} = \left(\frac{1-K}{K}\right)E_{vac}.$$

Dielectric breakdown occurs when the electric field within the dielectric becomes too strong. The solid body of the dielectric is shot through with microscopic cracks and voids. Passage of an ionizing particle from cosmic rays or background radiation ejects an electron from the wall of the void. This electron accelerates up the electric field lines crossing the void, gaining kinetic energy as it does so. Eventually, the electron will impact the wall of the void and deposit this kinetic energy in the form of heat, or if great enough, in the form of secondarily ejected electrons. A conducting path opens through the dielectric and a spark discharge takes place between the positive and negative sides of the dielectric.

The current along the path taken by the spark is frequently intense enough to damage or destroy the dielectric by burning, chemical decomposition, or melting. Brittle substances like glass or lucite crack under the thermal and electrical stresses.

#### 4.4.1.2 Gaseous Arc Discharge

Although the gas as a whole is overwhelmingly neutral, a small number of positive and negative ions always exist. Cosmic rays and nuclear background radiation, always present, are capable of ionizing molecules of the gas by ejecting one or more atomic electrons. The ejected electrons are accelerated up the electric field lines while the positive ions are acclereated down the field lines, both gaining kinetic energy according to

$$E_K = \tfrac{1}{2}mv^2 = eEd$$

where $d$ is the distance traveled. If the electric is sufficiently strong, or if the mean free path (the average distance traveled by a particle between collisions) is sufficiently long, then the ion or electron will acquire enough kinetic energy to ionize the next particle it hits. The original ion and the daughter ion are then accelerated once again with each creating a new ion at next impact. What begins as a single ion quickly becomes two, then four, etc., the total number doubling with each collision. The current through the gas quickly diverges. This is the phenomenon known as gaseous arc discharge.

# Chapter 5
# The Radiation Environment

## 5.2 Basic Radiation Physics

Becquerel discovered natural radioactivity in 1896 when he observed that photographic plates unexposed to light nevertheless "fogged" when left in the presence of uranium salts. This was shortly after Röentgen's discovery of x-rays and their ability to penetrate optically opaque matter and fog similar photographic plates. Becquerel immediately suspected that some similar undiscovered radiation was at work in his situation and in short order was able to demonstrate that the radiation in question emanated from the uranium salts. At first, he suspected that the uranium salts fluoresced x-rays when exposed to light but experiment showed that the radiation was constant in intensity with no discernable connection with the exposure of the salts to light. Further experiments showed that the "uranium emanations" as they were called, were actually three different types of radiation. One type of radiation was deflected in a magnetic field in a manner consistent with the presence of positive electric charge and was designated "$\alpha$ rays." A second type of radiation deflected in a magnetic field in a manner consistent with negative electric charge and was designated "$\beta$ rays." A third was undeflected by a magnetic field and was designated "$\gamma$ rays."

$\alpha$ rays were later shown to be the nuclei of helium atoms, consisting of two protons in conjunction with two neutrons. $\beta$ rays were shown to be electrons (although now the term is also used to designate the anti-matter equivalent of the electron, the positive electron, or positron), and $\gamma$ rays were shown to be very high energy electromagnetic radiation. Natural radioactivity turned out to be the spontaneous decay of the heaviest elements (uranium, thorium, actinium, etc.) into lighter elements through the emission of energetic particles. An atom of $^{238}$U , for example, decays into $^{234}$Th by the emission of one $\alpha$ particle. $^{234}$Th decays by $\beta$ particle emission into $^{234}$Pa . A sequence of $\alpha$ and $\beta$ particle emissions ensues culminating in a stable atom of lead: $^{206}$Pb . $\alpha$ particle emission in every case lowers the atomic number by two and the atomic weight by four. $\beta$ particle emission raises the atomic number by one and leaves the atomic weight unchanged. $\gamma$ ray emission occurs when an excited nucleus decays to the lowest energy state ("ground state") with no change in either atomic number or atomic weight.

# Chapter 6
# The Micrometeoroid/Orbital Debris Environment

## 6.1 Overview

The micrometeoroid population in low earth environment comes from sources in interplanetary or interstellar space, since the earth does not possess ring systems similar to those of Jupiter, Saturn, Uranus, and Neptune. Such ring systems are transient and short-lived on an astronomical timescale unless stabilized by the presence of "shepherd moons." The earth's moon is unfortunately too large in relation to its primary, therefore all matter left over from the formation of the earth and moon was long ago either accreted by the earth or moon, or swept out of earth orbit by chaotic gravitational interactions with the earth and moon.

Most micrometeoroids come from the breakup of comets. Comets are believed to originate in the Oort cloud, a reservoir of comets in roughly circular orbits at 20,000-100,000 astronomical units from the sun. The comets themselves are postulated to be condensed matter dating from the formation of the planets. They consist of dust, soot, and fine particles in a frozen matrix of ice, solid carbon dioxide, and other volatiles. In the Oort cloud, they are far enough from the sun to respond to gravitational perturbations from nearby stars which send a small but nearly constant number of them into orbits which bring them near the sun. Near perihelion, solar radiation vaporizes the volatile material within the comet, forming the long luminescent tail which makes the comet visible from earth. As the volatile matrix sublimates, the solid material is freed and the individual particles assume independent orbits roughly identical to that of the original comet.

Comets last for only about a thousand perihelion passages. Then the volatile material is all gone. Only the gritty residue remains. Gravitational perturbations first spread it throughout the original cometary orbit and then eventually throughout the entire volume of the solar system between the original comet's perihelion and aphelion. The earth sweeps it up as it passes through and it then becomes part of our micrometeoroid environment. Cometary orbits are not confined to the plane of the ecliptic, but may have any inclination and ascending node. As a result, micrometeroids of cometary origin may come at the earth from any direction in space and with a wide distribution of closing speeds.

The residue of active or recently deceased comets is still confined to a region of space near that of the comet itself. Unusually large meteor activity occurs when the earth passes through or near the orbit of a comet shortly before or after the comet itself passes by. Table 6.1 lists several prominent meteor showers, their recurring dates, and the comets with which they have been associated. Spacecraft vulnerability to micrometeoriod impact and damage rises proportionately with the increased flux at these times.

Cometary debris is small enough to be significantly affected by solar radiation pressure. Electromagnetic radiation carries momentum according to the relation

$$p = E/c,$$

so that the absorption or reflection of radiation delivers an impulse and therefore exerts a force. The pressure exerted by radiation is just the momentum flux delivered (the impulse per unit area per second), which is given by

$$P = \frac{p}{A\Delta t} = \frac{E}{cA\Delta t}.$$

$E/A\Delta t$ is the energy flux (Joules per square meter per second) of solar radiation and is known to be

$$\frac{E}{A\Delta t} = \left(\frac{R_s}{d}\right)^2 \sigma T_s^4$$

where $R_s$ is the radius of the sun and $d$ is distance from the sun. Assuming that the debris particles are spheres with cross-sectional area $A = \pi r^2$, the radiation force on a debris particle is

$$F_R = \frac{\pi \sigma r^2 R_s^2 T_s^4}{cd^2}.$$

But the solar gravitational force is

$$F_G = GM_s(4\pi r^3\rho/3)/d^2$$

where $\rho$ is the density of the particle. The ratio of the radiation force to the gravitational force is

$$\frac{F_R}{F_G} = \frac{3\sigma r_s^2 T_s^4}{4cGM_s\rho r} = \frac{5.78}{\rho r}$$

if $r$ is in m and $\rho$ is in kg m$^{-3}$. This expression is independent of distance $d$ from the sun, so that for particles of $\rho r < 5.78$, the radiation force exceeds gravitational force *everywhere*. Ordinary matter has a density $\rho \approx$1—10; this gives a critical particle radius of $r \approx 0.1$—$1\times10^{-6}$ m. Dust particles smaller than this size will be blown out of the solar system. As the radiation force effectively cancels the gravitational force, the particles will drift out of the solar system at their orbital speed and be gone in a matter of centuries.

For larger particles, $F_G >> F_R$, but radiation pressure can not be neglected. The orbital motion of the particle will make the radiation coming radially outward from the sun appear to come from a direction slightly in front of the particle. The component of momentum parallel to the the particle's velocity will impart a negative impulse to the particle and create a drag force of magnitude $(v/c)F_r$ that causes the particle to lose angular momentum and eventually fall into the sun. This is called the *Poynting-Robertson effect* and calculations indicate that a particle at a distance $d$ AU from the sun will fall into the sun in $t = 7 \times 10^6 \rho r d^2$ years. At a distance of 1 AU, a particle of $r = 10^{-6}$ m and $\rho = 4300$ kg m$^{-3}$ will fall into the sun in only 3000 years. Because of these two effects, supplemented by the solar wind but retarded slightly by planetary perturbations, micrometeoroids must be constantly replenished by the decay of contemporary comets and the introduction of new comets from the Oort cloud.

Larger meteoroids, including those large enough to survive entry into the earth's atmosphere and reach the earth's surface, are believed to be debris from asteroid collisions. On a rare but recurring basis, the earth has been directly hit by wayward asteroids and the occasional comet. The Tunguska strike of 1909 is believed to have been a comet which exploded at altitude in a fair imitation of a nuclear airburst without the radioactivity. The Barringer Meteor Crater in Arizona is the result of a middling-sized meteoroid impact. And the extinction of the dinosaurs at the end of the Cretaceous period is widely believed to have been the result of a large asteroid impact at Chincxulub, Mexico, on the Yucatan peninsula.

While such large impacts pose little risk to spacecraft because of their extremely low probability, they pose a risk to humanity that cannot be ignored. Some authorities estimate that deadly impacts of the size of the Tunguska event may occur as often as once every 1200 years. Should such an impact occur in populated areas, the loss of life and property would dwarf anything seen since Hiroshima and Nagasaki. Serious consideration has been given to establishing a program to locate, catalog, and track all asteroids large enough to cause such damage, and to develop methods for deflecting them when earth impact is predicted.

### 6.2 Hypervelocity Impact Physics

The damage done by micrometeoroid/orbital debris impact is a function of the kinetic energy possessed by the impacting particle. This in turn, goes up as the square of the closing velocity of the impacting particle and its target. The orbital velocity of the earth about the sun is 30 km/s. The maximum orbital velocity of a micrometeoroid at the earth's orbit is 42 km/s, solar system escape velocity at the radius of the earth's orbit. A micrometeoroid in retrograde orbit will meet the earth with a closing speed of 72 km/sec, while a micrometeoroid in prograde orbit will catch up with the earth at a closing speed of 12 km/s. The gravitational binding energy of the earth and micrometeoroid will contribute additional velocity according to

$$v_f^2 = v_i^2 + v_{esc}^2$$

where $v_i$ is the closing velocity of the micrometeoroid and the earth, $v_{esc}$ is the escape velocity of the earth, and $v_f$ is the final impact velocity of the micrometeoroid with the earth. This would raise the closing velocity of retrograde micrometeoroids to 72.8 km/s, and that of prograde micrometeoroids to 16 km/s. It also establishes a minimum impact velocity of 11 km/s for those micrometeoroids initially in orbits which pace the earth and have close to zero relative velocity.

Spacecraft impact velocities will be modified further by the spacecraft orbital motion. Orbital velocities in low earth orbit are on the order of 8 km/s. Head-on impact of a satellite in LEO with a micrometeoroid traveling at 72 km/s would yield a closing speed of 80 km/s. Impact from behind by the same micrometeoroid would take place at 64 km/s. Minimum velocity micrometeoroids would impact spacecraft at 19 km/s head-on and 3 km/s from behind. The full range of posssible micrometeoroid impact velocities for LEO satellites is therefore 3—80 km/s.

Orbital debris is by definition gravitationally bound to the earth and therefore traveling at speeds of 11 km/s or less. Spacecraft in LEO circular orbit traveling at 8 km/s are vulnerable to impacts at velocities ranging from head-on at 19 km/s to impact from behind at closing velocities of 3 km/s or less.

Impacts at 3—5 km/s or higher occur at speeds exceeding the speed of sound in both the impactor and the target. The impact generates shock waves propagating forward into the target and backward into the impactor. It is these shock waves which do the physical damage to both the impactor and target. The extent of the damage is a complicated function of both the speed of impact and the thickness of the target. If the thickness of the target $t$ is small compared to the diameter $d$ of the impactor, assumed spherical, i.e. $t/d << 1$, then the shock that propagates into the target can reflect off of the rear surface and overtake the shock propagating through the impactor. The reflected shock will interfere with and diminish the compressive shock in the impactor so that the impactor does not have time to get very hot and suffers very little damage. The impactor will proceed forward through the target intact. The target itself will either shatter or have a hole the shape and size of the impactor cut through it. As the thickness of the target relative to the diameter of the impactor increases, a critical value is reached where the compression wave in the impactor completely crosses the particle by the time the reflected wave catches up with it. Here, the compression wave has the opportunity to do substantial heating of the impactor and the particle melts or vaporizes. The interaction of the initial compression wave and the reflected wave impart an expansion to the particle material: it explodes. Both the impactor and target are destroyed. Above the critical value, $t/d >> 1$, the reflected wave will not have time to overtake the compression wave. The impactor will melt or vaporize, but the material of the particle will continue to move forward in cohesion. If the compression wave propagating forward through the target reaches the rear wall with sufficient strength, it will throw material from the back of the target and leave a crater. (See Figure 6.3 of Tribble.) For sufficiently large target thickness, the target compression wave will die out before reaching the far wall, and the impactor will bury itself in the target leaving a pit at its point of entry. As impactor velocity increases, the value of $t/d$ at which one damage mechanism transitions into the next decreases.

The adverse effects of hypervelocity impact on a spacecraft range from minor to catastrophic and depend in detail upon the speed and size of the impactor, the physical nature and function of the spacecraft components hit, and their position ralative to other parts of the spacecraft. Table 6.2 shows potential failure modes as a result of debris impact.

Minimum damage with the least probability of spacecraft failure occurs when the impactor is so small that the $t/d >> 1$ condition of the previous paragraph is satisfied. In this case, the impactor will bury itself in the target leaving only a small pit to mark its arrival. This will typically occur on the outer surfaces of the spacecraft, predominantly those facing the ram direction, and will affect such things as paints, thermal insulation, radiators, solar panels, windows, objective lenses and mirrors. Optical surfaces face the greatest threat at this level of impact: surface degradation and erosion increase light scattering thereby both increasing the optical noise and decreasing the image quality. Non-imaging components such as windows and solar panels suffer a loss of transparency — windows merely become harder to see through, but solar panels lose efficiency. Paint erosion and thermal insulation blanket perforation both expose protected areas to the space environment and compromise thermal conduction and radiation. Erosion of thermal radiator surfaces changes emissivity and absorptivity with consequent effects on spacecraft thermal control. Impact shock causes loss of material far beyond the perimeter of the impact crater. Laminated or multi-layered materials may peel off in comparatively large chunks, while brittle materials sustain cracks that may propagate some length. Impacts also create localized plasmas which can cause or contribute to electrostatic discharge and electronic component disruption or failure. All of these impact damage effects are magnified when combined with atomic oxygen attack and/or ultraviolet exposure.

Near the critical value of $t/d$, the impactor causes extensive physical damage even though it may not penetrate the target. As the compressive shock sent ahead of the impactor reflects off of the rear wall of the

**TABLE 6.1** Potential Failure Modes of Subsystems as a Result of Debris Impact

| Probable Critical Types of Failure | Subsystems | | | | | |
|---|---|---|---|---|---|---|
| | Pressure Cabins | Tanks | Radiators | Windows | Electronics | Special Surfaces |
| Catastrophic Rupture | × | × | | × | | |
| Detached Spaling | × | × | × | | × | |
| Secondary Fractures | | | × | | × | |
| Leakage | × | × | × | | | |
| Shock Pulse | × | | | × | × | |
| Vapor Flash | × | | | | | |
| Deflagration | | × | | | | |
| Deformation | | | × | | × | |
| Reduced Residual Strength | × | × | × | × | | |
| Fluid Contamination | | × | × | | | |
| Thermal Insulation Damage | × | × | | | | |
| Obscuration | | | | × | | |
| Erosion | | | | × | | × |

×, Subsystem is vulnerable to damage mode.
SOURCE: NASA, 1970.

target, it ejects small particles from that wall in a process called spallation. The spallated particles travel nearly as fast as the original impactor and can cause serious damage to the interior of the spacecraft. As impactor size increases to extreme values, $t/d << 1$, perforation takes place and the fragmented or liquified remains of the impactor and the target material it punches out damage an area much greater than the impact hole itself. The force of the impact causes buckling and bending of structural members, sending a shock wave traveling through the spacecraft fuselage and components.

If the spacecraft is crewed and therefore pressurized, the perforation allows a loss of cabin pressure. But even a perforation as large as one cm would cause a decompression slow enough to allow the crew a comfortable margin for plugging the leak and donning pressure suits. Far more serious is the light pulse emitted by flash vaporization of the impactor and wall material. This pulse can be bright enough to blind if looked at directly at the moment of impact. In addition, standard pressure for a crewed spacecraft is 14.7 psi, sea-level equivalent, which places the structural members of the crew module under considerable tension. Cracks in the fuselage beyond a critical length can lead to catastrophic fracture or uncontrolled crack propagation, a process known as "unzipping." Windows, being the weakest part of the fuselage, are the most susceptible to these types of failures. Unmanned spacecraft with pressurized components such as propellant tanks or cryogenic liquid storage tanks are also vulnerable to this type of destruction.

McKnight (1993) estimates that a collision with a ratio of impactor-kinetic-energy-to-target-mass exceeding 40 J/g will competely break up the target. The resulting fragments will disperse with a wide range of new velocities, placing them into a wide range of new orbits and adding to the orbital debris population. The number of new particles created is difficult to model and predict, but the quantity of particles created increases with decreasing size and could be in the millions for medium-sized particles.

### 6.4 The Orbital Debris Environment

Orbital debris is defined as anything artificial in orbit which isn't a functional spacecraft, either operational or on standby; see Figure 6.1. This definition covers everything from microscopic flecks of paint

to full-size spacecraft meters in length. Space objects fall into one of five categories: functional spacecraft, non-functional spacecraft, rocket bodies, mission-related debris, and fragmentation debris. Objects greater than 10 cm diameter can be reliably tracked and cataloged. The current census of cataloged space objects is about 9000 (Johnson, 1999). The growth of cataloged on-orbit objects by category and total is graphed in Figure 6.2. The proportion of cataloged space objects falling into each category is illustrated in Figure 6.3. Most spacecraft reside in low earth orbit (LEO) but there are also significant concentrations in circular semi-synchronous orbits, Molniya-type orbits, and in geosynchronous orbits (GEO). Orbital debris tends to be concentrated in these heavily used orbital regions but it is by no means confined to them.

Non-functional spacecraft and rocket bodies are important to the future evolution of the orbital debris environment because of their large size and potentially explosive propellants and batteries on board. More that 120 known breakups of non-functional spacecraft and rocket bodies have occurred releasing over 8000 cataloged pieces of fragmentation debris; more than 3000 of these debris objects remain in orbit.

Debris released in a breakup departs with a range of initial velocities and therefore with differing total gravitational energies. The semi-major axis of the orbit $a$ is related to the total energy $E$ by $a = -\frac{\mu}{2E}$ and the orbital period $T$ is related to the semi-major axis by $T = \frac{2\pi}{\sqrt{\mu}} a^{3/2}$. Because each item in the debris cloud has a different orbital period, the cloud disperses throughout the orbital plane. Then gravitational perturbations that depend on inclination and semi-major axis, such as regression of the nodes and rotation of the line of apsides (Bates, 1971), disperse the debris into all possible orbital planes. As a result, the debris cloud expands to fill a toroidal volume bounded by the perigee, apogee, and orbital inclination of the original object, as illustrated in Figure 6.4.

The products of slow deterioration from spacecraft aging and the effects of the space environment represent another large class of fragmentation debris. These fragments typically leave the spacecraft at low relative velocities and linger in the vicinity of the spacecraft for long periods. They occasionally are objects large enough to detect from Earth and catalog such as lost solar panels, pieces of debonded thermal blanket, or protective covers. The flaking of paint presents a more serious problem. In LEO the paint chips rapidly deorbit due to aerodynamic drag, but are just as rapidly replaced by new flaking.

Mission-related debris refers to objects deliberately released during the course of spacecraft deployment, activation, or operation. This category includes explosive bolts, spring release mechanicisms, spin-up and spin-down devices (e.g. yo-yos), stowage cords and protective covers. A detailed study of one Russian launch counted 76 objects released in space from either the launch vehicle or the spacecraft. Over 200 pieces of cataloged mission-related debris have been linked to the Mir space station, the majority intentionally dumped refuse. Since only the largest objects could be detected and cataloged, it may be safely concluded that the total number of items dumped overboard is much greater.

Solid rocket motor burns produce another type of mission-related debris in the form of aluminum oxide pellets. While probably no larger than 10 microns in diameter, as many as $10^{20}$ pellets may be produced in a single burn. Because of their small mass, the orbital lifetime of these particles is quite short, but the large number of solid rocket burns performed each year keeps the current population at significant levels as indicated by examination of impacts on exposed spacecraft surfaces.

The lifetime of orbital debris below 800 km is determined primarily by aerodynamic drag. The magnitude of the drag acceleration is proportional to the cross-sectional area per unit mass ($\beta = A/m$) and local atmospheric density $\rho$. $\rho$ falls off exponentially and is therefore greatest at the lower altitudes. It also varies with the solar cycle so that orbital decay is accelerated during times of solar maximum. During the last two solar maxima, rate of orbital decay exceeded rate of space object generation during new launches so that the total number of cataloged objects actually declined. Figure 6.5 shows the dependence of orbital lifetime on altitude, solar cycle phase, and spacecraft $A/m$.

At higher altitudes solar radiation and solar and lunar perturbations predominate. Their effects are greatest on highly elliptical orbits such as geosynchronous transfer orbits (GTO). Solar and lunar perturbations depend critically on alignment of the sun, moon, and orbital plane. Some configurations accelerate orbital decay while others inhibit it. As a result, GTO orbital decay could take anywhere from a few months to a few centuries. Solar radiation pressure is only effective on objects with a large area-to-mass ratio. The net effect of solar radiation is to increase the eccentricity of the orbit while simultaneously lowering perigee. Orbital decay is accelerated if the perigee is lowered enough to significantly enhance aerodynamic drag.

Over 16,000 cataloged objects have re-entered since 1957. Presently, two to three cataloged objects re-enter each day, primarily large objects launched into low orbits and small objects with large $\beta$. Large objects initially placed into orbits higher than 600 km seldom re-enter.

### 6.9 References


1. Bates, R., D. Meuller, J. White. 1971. Fundamentals of Astrodynamics. New York: Dover Publications, Inc. pp. 157-8.
2. Johnson, N., J. Loftus, Jr., 1999. Reducing Orbital Debris: Standards and Practices. Launchspace Magazine, March/April 1999. p. 24.
3. McKnight, D.S. 1993. Collision and Breakup Models: Pedigree, Regimes, and Validation/Verification. Briefing presented to the National Research Council Committee on Space Debris Workshop, Irving, California, November 18.
4. NASA (National Aeronautics and Space Administration) 1970. Meteoroid Damage Assessment. NASA Space Vehicle Design Criteria. NASA-SP8042. Washington, D.C.: NASA.